The Journal of Antibiotics (2015) 68, 385–394

www.nature.com/ja

ORIGINAL ARTICLE

Novel terpenes generated by heterologous expression of bacterial terpene synthase genes in an engineered *Streptomyces* host

Yuuki Yamada[1], Shiho Arima[2], Tohru Nagamitsu[2], Kohei Johmoto[3], Hidehiro Uekusa[3], Tadashi Eguchi[3], Kazuo Shin-ya[4], David E Cane[5] and Haruo Ikeda[1]

Mining of bacterial genome data has revealed numerous presumptive terpene synthases. Heterologous expression of several putative terpene synthase genes in an engineered *Streptomyces* host has revealed 13 newly discovered terpenes whose GC-MS and NMR data did not match with any known compounds in spectroscopic databases. Each of the genes encoding the corresponding terpene synthases were silent in their parent microorganisms. Heterologous expression and detailed NMR spectroscopic analysis allowed assignment of the structures of 13 new cyclic terpenes. Among these newly identified compounds, two were found to be linear triquinane sesquiterpenes that have never previously been isolated from bacteria or any other source. The remaining 11 new compounds were shown to be diterpene hydrocarbons and alcohol, including hydropyrene (1), hydropyrenol (2), tsukubadiene (11) and odyverdienes A (12) and B (13) each displaying a novel diterpene skeleton that had not previously been reported.
The Journal of Antibiotics (2015) **68,** 385–394; doi:10.1038/ja.2014.171; published online 21 January 2015

INTRODUCTION

Among the tens of thousands of known terpenoid metabolites, many with medically or agriculturally useful activity, the vast majority have been isolated from plants or fungi, with only a relative handful having been obtained from bacterial sources. Among the bacteria, *Actinomycetales* microorganisms have been known for many years as producers of volatile odoriferous metabolites, foremost among these the degraded sesquiterpene alcohol geosmin that is largely responsible for the characteristic odor of moist soil.[1,2] Also contributing to this odor is the well-recognizable musty scent of the methylated monoterpene alcohol 2-methylisobornenol also produced by many *Actinomycetales* microorganisms.[3] These two terpene natural products have also been detected from *Cyanobacteria*[4,5] and *Myxobacteria*.[6]

The numerous parent cyclic monoterpene, sesquiterpene and diterpene hydrocarbons and alcohols are all formed by variations in a universal cyclization mechanism that is initiated by an enzymatic ionization of the relevant acyclic allylic diphosphate precursors, geranyl diphosphate, farnesyl diphosphate and geranylgeranyl diphosphate, respectively. Intramolecular attack by the resultant allylic cations on the remaining double bonds, followed by a cascade of electrophilic reactions, frequently in combination with common carbocation rearrangements, terminated by quenching of the positive charge by deprotonation or by capture of water generates an enormous variety of cyclic terpenes, with nearly 400 parent skeletons being known to date. The experimental or bioinformatics search for the responsible bacterial terpene synthases has been hindered since these microbial terpene synthases exhibit at most low and most often insignificant levels of overall amino acid sequence similarity to those from fungi and plants as well as relatively low levels of mutual sequence similarity even among the bacterial terpene synthases. To address this challenge, we have developed a powerful genome mining method using hidden Markov models to carry out Pfam searches of bacterial sequence databases. As originally reported, the initial formulation of this model resulted in the recognition of 41 bacterial terpene synthases, including for the first time those for 2-methylisoborneol and 2-methylenebornane.[7] Although the original PF03936 profile (terpene synthase C-terminal domain) had been based exclusively on an alignment of plant terpene synthase sequences, we subsequently generated a new hidden Markov model profile for searching for bacterial terpene synthases by using the alignment data derived from the 41 bacterial terpene synthase sequences that had been discovered by the first round of hidden Markov model searching.[8] Most recently, we have used the 140 bacterial terpene synthase sequences uncovered in the second round of Pfam searching to create a third generation hidden Markov model, thereby uncovering 262 prospective bacterial terpene synthases by screening of >8.6 million predicted protein sequences found within the most recent public and in-house bacterial genome databases.[9] These results establish that

[1]Kitasato Institute for Life Sciences, Kitasato University, Kitasato, Sagamihara, Kanagawa, Japan; [2]School of Pharmacy, Kitasato University, Shirokane, Minato-ku, Tokyo, Japan; [3]Department of Chemistry and Materials Science, Tokyo Institute of Technology, O-okayama, Meguro-ku, Tokyo, Japan; [4]National Institute of Advanced Industrial Science and Technology, Aomi, Koto-ku, Tokyo, Japan and [5]Department of Chemistry, Box H, Brown University, Providence, RI, USA
Correspondence: Professor H Ikeda, Kitasato Institute for Life Sciences, Kitasato University, 1-15-1 Kitasato, Minami-ku, Sagamihara, Kanagawa 252-0373, Japan.
E-mail: ikeda@ls.kitasato-u.ac.jp
Received 25 September 2014; revised 20 November 2014; accepted 12 December 2014; published online 21 January 2015

terpene synthases are widely distributed in bacteria. Although many of the responsible genes appear to be silent in their host microorganisms, at least under common laboratory culture conditions, we have been able to express a large number of bacterial genes encoding terpene synthases in an engineered *Streptomyces* host, thereby enabling the identification of many terpenoid metabolites from this heterologous expression system. Although the majority of the newly detected bacterial terpenes corresponded to known compounds previously isolated from fungi or plants, several novel terpenes were also isolated.[9]

We now report the detailed structure elucidation of more than a dozen novel terpenes generated by heterologous expression of the newly identified *Streptomyces* genes that encode terpene synthases.

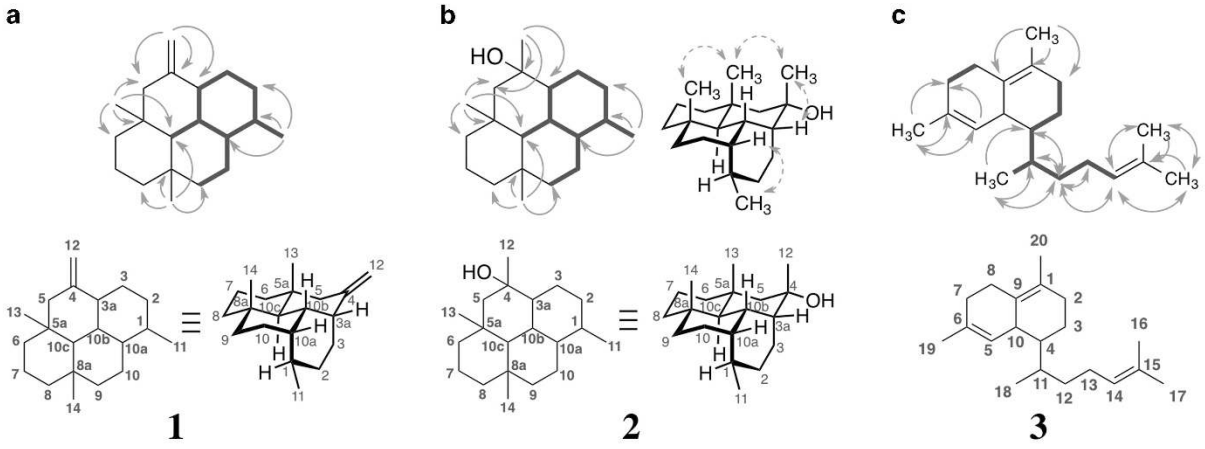
RESULTS

Diterpenes from transformants carrying *sclav_p0765*

Cultures of the large-deletion mutant, *S. avermitilis* SUKA22, carrying *sclav_p0765* from *S. clavuligerus* ATCC 26074, known as a producer of the clinically important antibiotics, clavulanic acid and cephamycin C, produced ten diterpenes (Supplementary Figure S1). Although six of these were obtained in only very-low quantities, the four major compounds were purified. After cultivation of the transformants in 35 liters (5 × 7 liters), the terpene products were extracted from the mycelium and the four major components were purified by silica gel column chromatography. The HR-MS (electron ionization; EI) of peak 5 (hydropyrene; **1,** 51.7 mg, $[\alpha]_D^{24}$ −60.7 (*c* 0.1, $CHCl_3$), IR ν_{max} (attenuated total reflection) 2921, 1440, 1382 and 890 cm^{-1}) showed a molecular ion peak at *m/z* 272.2505 $[M]^+$, consistent with a molecular formula of $C_{20}H_{32}$ (calculated 272.2504) for a diterpene hydrocarbon with five degrees of unsaturation. The ^1H NMR spectrum (Supplementary Figure S8) of **1** showed two olefinic protons (δ_H 4.42, d, J = 2.0 Hz, 1 H, H-12; 4.65, t, J = 2.0 Hz, 1 H, H-12; Figure 1a) assigned to an exomethylene. Analysis of the ^{13}C NMR (Supplementary Figure S8) and DEPT spectra of **1** showed 20 resolved signals and confirmed the presence of three methyls, eight sp^3 methylenes, one sp^2 methylene, five sp^3 methines and one sp^2 and two sp^3 quaternary carbons. One double bond accounted for one degree of unsaturation, whereas the remaining degrees of unsaturation were thus attributed to four rings in **1**. The proton and carbon connectivity were assigned by analysis of the ^1H–^{13}C HMQC spectrum (Supplementary Figure S9), as shown in Figure 1a. Analysis of the ^1H–^1H COSY (Supplementary Figure S10) revealed one partial structure C-1– C-11 (Figure 1a) and the ^1H–^{13}C long-range couplings of *2J* and *3J* observed in the ^1H–^{13}C HMBC experiments (Supplementary Figure S11) gave the following information: the cross peaks from H_3-11 (δ_H 0.83) to C-1 (δ_C 28.2), C-2 (δ_C 35.9) and C-10a (δ_C 44.0), from H_2-12 (δ_H 4.42, 4.65) to C-3a (δ_C 47.6), C-4 (δ_C151.8) and C-5 (δ_C 49.8), from H_3-13 (δ_H 0.77) to C-5, C-5a (δ_C 36.2), C-6 (δ_C 42.4) and C-10c (δ_C 45.8) and from H_3-14 (δ_H 0.90) to C-8 (δ_C 42.3), C-8a (δ_C 34.6), C-9 (δ_C 39.8) and C-10c supported the structure as shown in Figure 1a.

The molecular formula of peak 10 (hydropyrenol; **2,** 22.3 mg, $[\alpha]_D^{24}$ -124.7 (*c* 0.1, $CHCl_3$), IR ν_{max} (attenuated total reflection) 3434, 2921, 1440, 1382 and 890 cm^{-1}) was deduced as $C_{20}H_{32}O$ (calculated 290.2610) by HR-MS (EI) *m/z* 290.2612 $[M^+]$ for a diterpene alcohol with four degrees of unsaturation. The ^1H and ^{13}C NMR spectra (Supplementary Figure S13) of **2** were similar to those of **1**. Analysis of ^1H–^1H COSY (Supplementary Figure S14) revealed one partial structure C-1 to C-11 (Figure 1b), ^1H–^{13}C HMQC (Supplementary Figure S15) and ^1H–^{13}C HMBC (Supplementary Figure S16) revealed that **2** is the C-4 alcohol derivative of **1** (Figure 1b). The ^1H–^{13}C long-range couplings of *2J* and *3J* observed in the ^1H–^{13}C HMBC experiments gave the following information: the cross peaks from H_3-11 (δ_H 0.82) to C-1 (δ_C 28.3), C-2 (δ_C 36.1) and C-10a (δ_C 44.1), from H_3-12 (δ_H 1.08) to C-3a (δ_C 49.9), C-4 (δ_C 73.8) and C-5 (δ_C 52.3), from H_3-13 (δ_H 1.13) to C-5, C-5a (δ_C 34.9), C-6 (δ_C 43.2) and C-10c (δ_C 45.2) and from H_3-14 (δ_H 0.99) to C-8 (δ_C 42.6), C-8a (δ_C 34.6), C-9 (δ_C 39.8) and C-10c supported the structure as shown in Figure 1b. The stereochemical assignments of the methyl protons (H_3-12) were elucidated from the ^1H–^1H NOESY data (Supplementary Figure S17). Two pairs of correlations between two protons H_3-12/H_1-3a and H_3-12/H_3-13, respectively, were observed by ^1H–^1H NOESY, suggesting that the hydroxyl residue at C-4 is equatorial (Figure 1b). Since NMR analysis was not sufficient to elucidate the full structures of **1** and **2**, X-ray crystallographic analysis of **1** was undertaken. Although neither **1** nor **2** formed fine crystals in variety of solvent system, the corresponding 4,12-epoxide of **1** (**1-epoxide**; ^1H NMR ($CDCl_3$, 400 MHz) δ_H 2.70 (1H, d, J = 5.0 Hz), 2.69 (1H, d, J = 5.0 Hz), 2.04 (1H, dt, J = 12.8, 5.0 Hz), 1.93 (1H, d, J = 12.8 Hz), 1.79–0.83 (18H, m), 0.96 (3H, s), 0.95 (3H, s), 0.85 (3H, d, J = 6.2 Hz), 0.64 (1H, dd, J = 12.8, 1.4 Hz)) prepared by oxidation of the C-4,12 olefinic bond formed tiny monoclinic crystals (Supplementary Figure S19). Crystal data for **1-epoxide**: empirical formula; $C_{20}H_{32}O$; molecular weight: 288.46; crystal system: monoclinic; space group: $P2_1$; unit cell dimensions: a = 6.2239(1) Å, b = 8.9584(2) Å, c = 14.6575(3) Å, β = 93.696(1)°, V = 815.55(3) $Å^3$; *Z*-value = 2; D_{calc} = 1.175 g cm^{-3}; T = 173(2) K, no. of unique reflections = 2924, R_{int} = 0.0357, no. of parameters = 193, no. of restraints = 1, R_1 = 0.0398, wR_2 = 0.1091, S = 1.091 for 2402 reflections, max/min. Residual density 0.16/-0.15 $eÅ^{-3}$. The absolute structure was unable to be determined from Flack parameter 0.1(2). Diterpenes, **1** and **2** each, have novel 6-6-6-6 ring skeletons.

The compound corresponding to peak 2 purified (8.0 mg) was identical to the known diterpene isoelisabethatriene[14] (isolated from sea plumes *Pseudopterogorgia elisabethae*). The ^1H and ^{13}C NMR data of peak 3 (isoelisabethatriene B, **3**, 4.8 mg, $[\alpha]_D^{24}$ +9.9 (*c* 0.1, $CHCl_3$), IR ν_{max} (attenuated total reflection) 2921, 1440, 1382 and 890 cm^{-1}) were very similar to those of isoelisabethatriene. The HR-MS (EI) of **3** showed a molecular ion peak at *m/z* 272.2505 $[M]^+$ consistent with $C_{20}H_{32}$ (calculated 272.2504) for a diterpene hydrocarbon with five degrees of unsaturation. The ^1H NMR spectrum (Supplementary Figure S20) of **3** showed two olefinic protons (δ_H 5.11, t, J = 7.0 Hz, 1H, H-14; 5.40, brs, 1H, H-5; Figure 1c). Analysis of ^{13}C NMR (Supplementary Figure S20) and DEPT spectra of **3** showed 20 resolved signals and confirmed the presence of five methyls, six sp^3 methylenes, three sp^3 methines, two sp^2 methines and four sp^2 quaternary carbons. Three double bonds accounted for three degrees of unsaturation, indicating a bicyclic skeleton for **3**. The connectivity of proton and carbon were assigned by analysis of the ^1H–^{13}C HMQC spectrum (Supplementary Figure S21) which combined with analysis of the ^1H–^1H COSY (Supplementary Figure S22) and ^1H–^{13}C HMBC (Supplementary Fig. S23) spectra. The ^1H–^1H COSY revealed two partial structures C-2–C-14 and C-7–C-8 (Figure 1c). The ^1H–^{13}C long-range couplings of *2J* and *3J* observed in the ^1H–^{13}C HMBC experiments gave the following information: the cross peaks from H_1-4 (δ_H 1.16) to C-12 (δ_C 35.8), from H_1-5 (δ_H 5.40) to C-7 (δ_C 31.9) and C-19 (δ_C 23.5), from H_1-11 (δ_H 1.24) to C-12, from H_2-12 (δ_H 1.30) to C-4 (δ_C 43.2), C-11 (δ_C 31.5), C-13 (δ_C 26.2), C-14 (δ_C 124.9) and C-18 (δ_C 13.9), from H_2-13 (δ_H 2.01, 1.96) to C-12, from H_1-14 (δ_H 5.11) to C-12, C-16 (δ_C 18.0) and C-17 (δ_C 25.7), from H_3-16 (δ_H 1.60) to C-14, C-15 (δ_C 131.2) and C-17, from H_3-17 (δ_H 1.69) to C-14, C-15 and C-16, from H_3-18 (δ_H 0.77) to C-4, C-11 and



NMR chemical shifts in $CDCl_3$

	1		2		3		
No.	δ_C	δ_H (multi, *J* in Hz)	δ_C	δ_H (multi, *J* in Hz)	No.	δ_C	δ_H (multi, *J* in Hz)
1	28.2	1.66 (1H, m)	28.3	1.67 (1H, m)	1	130	-
2	35.9	1.02 (1H, m), 1.66 (1H, m)	36.1	0.97 (1H, m), 1.68 (1H, m)	2	32.4	1.92 (1H, m), 1.89 (1H, m)
3	29.6	1.26 (1H, m), 1.74 (1H, m)	25.1	1.26 (1H, m), 1.74 (1H, m)	3	21.5	1.53 (1H, m), 1.17 (1H, m)
3a	47.6	2.27 (1H, m)	49.9	1.47 (1H, m)	4	43.2	1.16 (1H, m)
4	151.8	-	73.8	-	5	124.4	5.40 (1H, brs)
5	49.8	1.50 (1H, m), 2.26 (1H, m)	52.3	1.14 (1H, m), 1.33 (1H, m)	6	134.2	-
5a	36.2	-	34.9	-	7	31.9	1.98 (1H, m), 1.83 (1H, m)
6	42.4	1.17 (1H, m), 1.35 (1H, m)	43.2	1.02 (2H, m)	8	26.7	2.70 (2H, ddd, *J* = 12.9, 4.5, 2.5)
7	18.9	1.44 (2H, m)	18.2	1.41 (2H, m)	9	124.5	-
8	42.3	1.15 (1H, m), 1.33 (1H, m)	42.6	1.11 (1H, m), 1.36 (1H, m)	10	39.2	2.53 (1H, brs)
8a	34.6	-	34.6	-	11	31.5	1.24 (1H, m)
9	39.8	0.91 (1H, m), 1.24 (1H, m)	39.8	0.92 (1H, m), 1.29 (1H, m)	12	35.8	1.30 (2H, m)
10	22.7	1.51 (1H, m), 1.63 (1H, m)	22.8	1.51 (1H, m), 1.63 (1H, m)	13	26.2	2.01 (1H, m), 1.96 (1H, m)
10a	44.0	1.23 (1H, m)	44.1	1.25 (1H, m)	14	124.9	5.11 (1H, brt, *J* = 7.0)
10b	36.1	1.89 (1H, m)	31.6	2.17 (1H, m)	15	131.2	-
10c	45.8	1.37 (1H, m)	45.2	1.25 (1H, m)	16	18.0	1.60 (3H, s)
11	20.7	0.83 (3H, d, *J* = 7.6)	20.6	0.82 (3H, d, *J* = 7.6)	17	25.7	1.69 (3H, s)
12	106.7	4.42 (1H, d, *J* = 2.0), 4.65 (1H, t, *J* = 2.0)	31.1	1.08 (3H, s)	18	13.9	0.77 (3H, d, *J* = 6.8)
					19	23.5	1.66 (3H, s)
13	20.4	0.77 (3H, s)	24.1	1.13 (3H, s)	20	18.5	1.64 (3H, s)
14	21.0	0.90 (3H, s)	22.0	0.99 (3H, s)			

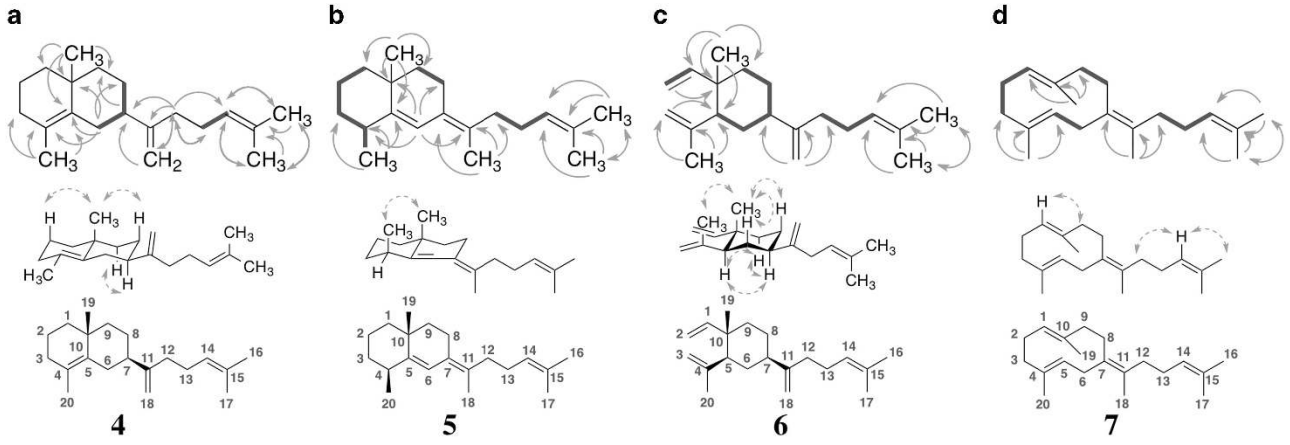
Figure 1 Structures of hydropyrene ((**a**); **1**), hydropyrenol ((**b**); **2**) and isoelisabethatriene B ((**c**); **3**) produced by *S. avermitilis* SUKA22 carrying *sclav_p0765*. The upper panel shows ^{1}H–^{1}H COSY (bold line), ^{1}H–^{13}C HMBC (arrow) and ^{1}H–^{1}H NOESY (dashed double arrow) data for each compound. Double arrows indicate a cross peak in HMBC. A full color version of this figure is available at *The Journal of Antibiotics* journal online.

C-12, from H_3-19 (δ_H 1.66) to C-5, C-6 (δ_C 134.2) and C-7 and from H_3-20 (δ_H 1.64) to C-1 (δ_C 134.2), C-2 (δ_C 32.4) and C-9 (δ_C 124.5) supported the structure, a isomer of isoelisabethatriene, as shown in Figure 1c.

Diterpenes from transformants carrying *sclav_p1169*

S. avermitilis SUKA22 transformants carrying *sclav_p1169* of *S. clavuligerus* ATCC 26074 produced seven diterpene compounds (Supplementary Figure S2). These metabolites, which accumulated in the mycelium, were purified from a 7-liter culture by silica gel column chromatography. Six components, peak 4 (clavulatriene A; **4**, 22.3 mg, $[\alpha]_D^{24}$ +42.0 (*c* 0.15, $CHCl_3$), IR ν_{max} (attenuated total reflection) 2921, 1440, 1382 and 890 cm^{-1}), peak 7 (clavulatriene B; **5**, 4.2 mg, $[\alpha]_D^{24}$ −88.3 (*c* 0.1, $CHCl_3$), IR ν_{max} (attenuated total reflection) 2921, 1440, 1382 and 890 cm^{-1}, λ_{max}^{MeOH} nm (ε) 246 (1003)), peak 1 (prenyl-β-elemene, **6**; 0.9 mg, $[\alpha]_D^{24}$ +15.1 (*c* 0.1, $CHCl_3$), IR ν_{max} (attenuated total reflection) 2921, 1440, 1382 and 890 cm^{-1}), peak 2 (prenylgermacrene B; **7**, 1.3 mg, $[\alpha]_D^{24}$ −110.2 (*c* 0.1, $CHCl_3$), IR ν_{max} (attenuated total reflection) 2921, 1440, 1382 and 890 cm^{-1}), peak 3 (prenylgermacrene;[15] 1.2 mg), and peak 5 (lobophytumin C;[16] 4.8 mg), were isolated. Each of the latter two metabolites has previously been isolated from plant and soft coral sources, respectively.

The HR-MS (EI) of **4** showed a molecular ion peak at *m/z* 272.2506 $[M]^+$, corresponding to $C_{20}H_{32}$ (calculated. 272.2504) for a diterpene hydrocarbon with five degrees of unsaturation. The ^{1}H NMR spectrum (Supplementary Figure S25) of **4** showed two olefinic protons (δ_H 4.74, d, *J* = 1.6 Hz, 1H, H-18; 4.80, brs, 1H, H-18; Figure 2a) assignable to an exomethylene. An additional olefinic proton (δ_H 5.14, m, 1H, H-14) was also observed. Analysis of the ^{13}C NMR (Supplementary Figure S25) and DEPT spectra of **4** showed 20



NMR chemical shifts in $CDCl_3$

	4		5		6		7	
No.	δ_C	δ_H (multi, *J* in Hz)	δ_C	δ_H (multi, *J* in Hz)	δ_C	δ_H (multi, *J* in Hz)	δ_C	δ_H (multi, *J* in Hz)
1	42.4	1.26 (1H, m), 1.55 (1H, m)	41.7	1.15 (1H, m), 1.53 (1H, m)	150.1	5.81 (1H, dd, *J* = 10.7, 17.7)	128.4	4.71 (1H, m)
2	19.1	1.55 (2H, m)	17.5	1.47 (1H, m), 1.54 (1H, m)	109.7	4.91 (1H, d, *J* = 17.5), 4.89 (1H, s)	26.9	2.30 (1H, m), 2.08 (1H, m)
3	33.2	1.83 (2H, m)	33.2	1.55 (2H, m)	112.0	4.81 (1H, brs,), 4.59 (1H, d, *J* = 1.2)	39.0	2.18 (1H, m), 2.13 (1H, m)
4	124.5	-	38.2	2.50 (1H, m)	147.8	-	135.8	-
5	135.1	-	148.5	-	52.8	1.99 (1H, m)	124.6	4.51 (1H, m)
6	31.3	2.55 (1H, dt, *J* = 14.0, 2.4), 1.74 (1H, m)	120.8	6.12 (1H, s)	33.3	1.57 (1H, m), 1.53 (1H, m)	28.4	2.00 (2H, m)
7	45.6	1.80 (1H, m)	128.5	-	44.4	1.95 (1H, m)	124.7	-
8	28.2	1.54 (1H, m), 1.63 (1H, m)	22.9	2.20 (1H, m), 2.48 (1H, m)	26.8	1.62 (1H, m), 1.41 (1H, m)	39.8	1.90 - 2.00 (2H, m)
9	40.3	1.28 (1H, m), 1.50 (1H, m)	41.5	1.35 (1H, m), 1.45 (1H, m)	40.1	1.44 (1H, m), 1.41 (1H, m)	40.0	1.90 - 2.00 (2H, m)
10	34.5	-	34.6	-	40.0	-	131.4	-
11	154.8	-	128.4	-	154.5	-	131.7	-
12	35.0	2.07 (2H, m)	34.9	2.01 (1H, m), 2.15 (1H, m)	34.9	2.06 (1H, m), 2.03 (1H, m)	34.7	2.08 (1H, m), 2.02 (1H, m)
13	26.9	2.11 (2H, m)	26.7	2.02 (1H, m), 2.13 (1H, m)	27.3	2.11 (2H, m)	26.0	2.02 (2H, m)
14	124.3	5.14 (1H, m)	124.5	5.14 (1H, m)	124.2	5.13 (1H, m)	124.6	5.04 (1H, m)
15	131.5	-	131.7	-	131.6	-	131.6	-
16	25.7	1.69 (3H, s)	25.7	1.67 (3H, s)	25.7	1.68 (3H, s)	25.8	1.69 (3H, s)
17	17.7	1.62 (3H, s)	17.6	1.60 (3H, s)	17.7	1.61 (3H, s)	17.7	1.62 (3H, s)
18	106.8	4.74 (1H, d, *J* = 1.6), 4.80 (1H, brs)	17.7	1.78 (3H, d, *J* = 2.1)	107.0	4.79 (1H, brs), 4.72 (1H, brs)	18.0	1.68 (3H, s)
19	24.7	1.04 (3H, s)	26.2	1.13 (3H, s)	16.6	0.99 (3H, s)	16.4	1.58 (3H, s)
20	19.3	1.60 (3H, s)	23.3	1.15 (3H, d, *J* = 12.2)	24.7	1.69 (3H, s)	16.1	1.53 (3H, s)

Figure 2 Structures of clavulatriene (**a**) (A; **4**), clavulatriene (**b**) (B; **5**), prenyl-β-elemene ((**c**); **6**) and prenylgermacrene B ((**d**); **7**) produced by *S. avermitilis* SUKA22 carrying *sclav_p1169*. The upper panel shows ^{1}H–^{1}H COSY (bold line), ^{1}H–^{13}C HMBC (arrow) and ^{1}H–^{1}H NOESY (dashed arrow) data for each compound. Double arrows indicate a cross peak in ^{1}H–^{13}C HMBC. A full color version of this figure is available at *The Journal of Antibiotics* journal online.

resolved signals and confirmed the presence of four methyls, eight sp^3 methylenes, one sp^2 methylene, one sp^3 methine, one sp^2 methine and four sp^2 and one sp^3 quaternary carbons. Three double bonds accounted for three degrees of unsaturation, indicating that **4** possessed a bicyclic skeleton. The connectivity of proton and carbon were assigned by analysis of the ^{1}H–^{13}C HMQC spectrum (Supplementary Figure S26), as shown in Figure 2a. Analysis of the ^{1}H–^{1}H COSY (Supplementary Figure S27) revealed two partial structures C-6–C-8 and C-13–C-14 (Figure 2a) and the ^{1}H–^{13}C long-range couplings of *2J* and *3J* observed in the ^{1}H–^{13}C HMBC (Supplementary Figure S28) experiments gave the following information: the cross peaks from H_2-6 (δ_H 1.60) to C-4 (δ_C 124.5), C-5 (δ_C 135.1), C-8 (δ_C 28.2) and C-10 (δ_C 34.5), from H_1-7 (δ_H 1.80) to C-9 (δ_C 40.3), from H_2-12 (δ_H 2.07) to C-7 (δ_C 45.6), C-11 (δ_C 154.8), C-13 (δ_C 26.9), C-14 (δ_C 124.3) and C-18 (δ_C 106.8), from H_2-13 (δ_H 2.11) to C-12 (δ_C 35.0), from H_1-14 (δ_H 5.14) to C-16 (δ_C 25.7) and C-17 (δ_C 17.7), from H_3-16 (δ_H 1.69) to C-14, C-15 (δ_C 131.5) and C-17, from H_3-17 (δ_H 1.62) to C-14, C-15 and C-16, from H_3-19 (δ_H 1.04) to C-1 (δ_C 42.4), C-5 (δ_C 135.1), C-9 (δ_C 40.3) and C-10 (δ_C 34.5) and from H_3-20 (δ_H 1.60) to C-3 (δ_C 33.2) and C-4 and C-5 supported the structure as shown in Figure 2a. The stereochemical assignment of the methyl protons (H_3-19; two pairs of correlations between two protons H_3-19/H_2-2 and H_3-19/H_2-8, respectively) was elucidated from ^{1}H–^{1}H NOESY data (Figure 2a and

Supplementary Figure S29). Since a correlation between two protons H_1-7 and H_1-9 was observed by ^1H–^1H NOESY, the prenyl residue at C-7 was assigned an equatorial configuration (Figure 2a).

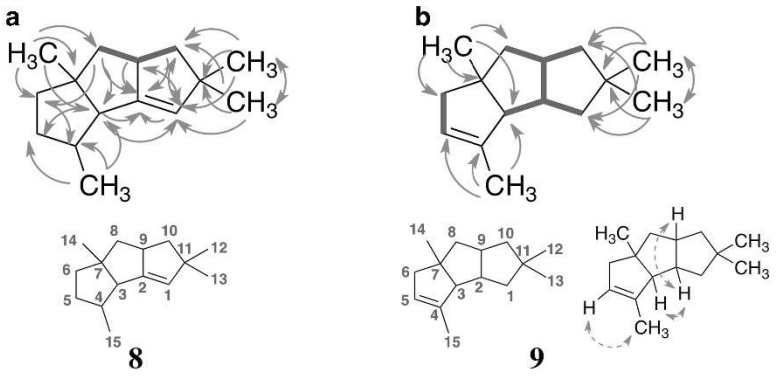
The HR-MS (EI) of **5** showed a molecular ion peak at *m/z* 272.2506 $[M]^+$, consistent with a molecular formula of $C_{20}H_{32}$ (calculated. 272.2504) for a diterpene hydrocarbon with five degrees of unsaturation. The ^1H NMR spectrum (Supplementary Figure S31) showed two olefinic protons (δ_H 5.14, m, 1H, H-14; 6.12, s, 1H, H-6; Figure 2b). The ^{13}C NMR (Supplementary Figure S31) and DEPT spectra of **5** showed 20 resolved signals and indicated the presence of five methyls, seven sp^3 methylenes, one sp^3 methine, two sp^2 methines, and four sp^2 and one sp^3 quaternary carbons. Three double bonds accounted for three degrees of unsaturation, with the remaining two degrees of unsaturation was attributed to a bicyclic skeleton in **5**. The connectivity of proton and carbon were assigned by analysis of the ^1H–^{13}C HMQC spectrum (Supplementary Fig. S32), as shown in Figure 2b. Analysis of the ^1H–^1H COSY (Supplementary Figure S33) revealed three partial structures C-1–C-20, C-8–C-9 and C-12–C-14 (Figure 2b) and the ^1H–^{13}C long-range couplings of *2J* and *3J* observed in the ^1H–^{13}C HMBC (Supplementary Figure S34) experiments gave the following information: the cross peaks from H_1-6 (δ_H 6.12) to C-4 (δ_C 38.2), C-8 (δ_C 22.9), C-10 (δ_C 34.6) and C-11 (δ_C 128.4), from H_3-16 (δ_H 1.67) to C-14 (δ_C 124.5), C-15 (δ_C 131.7) and C-17 (δ_C 17.6), from H_3-17 (δ_H 1.60) to C-14, C-15 and C-16 (δ_C 25.7), from H_3-18 (δ_H 1.78) to C-7 (δ_C 128.5), C-11 (δ_C 128.4) and C-12 (δ_C 34.9), from H_3-19 (δ_H 1.13) to C-1 (δ_C 41.7), C-5 (δ_C 148.5), C-9 (δ_C 41.5) and C-10 (δ_C 34.6), and from H_3-20 (δ_H 1.15) to C-3 (δ_C 33.2), C-4 and C-5 supported the structure as shown in Figure 2b, while and the stereochemistry of two methyl protons (H_3-19 and H_3-20) was deduced from the ^1H–^1H NOESY data (Figure 2b and Supplementary Figure S35).

The HR-MS (EI) of **6** showed a molecular ion peak at *m/z* 272.2506 $[M]^+$, consistent with a molecular formula of $C_{20}H_{32}$ (calculated 272.2504) for a diterpene hydrocarbon with five degrees of unsaturation. The ^1H NMR spectrum (Supplementary Figure S37) showed six olefinic protons (δ_H 4.59, d, $J = 1.2$ Hz, 1H, H-3; 4.72, brs, 1H, H-18; 4.79, brs, 1H, H-18; 4.81, brs, 1H, H-3; 4.89, s, 1H, H-2; 4.91, d, $J = 17.5$ Hz, 1H, H-2; Figure 2c) assignable to three pairs of exomethylene protons. Two additional olefinic protons (δ_H 5.13, m, 1H, H-14; 5.81, dd, $J = 10.7$, 17.7 Hz, 1H, H-1) were also observed. The ^{13}C NMR (Supplementary Figure S37) and DEPT spectra of **6** showed 20 resolved signals and indicated the presence of four methyls, five sp^3 methylenes, three sp^2 methylenes, two sp^3 methines, two sp^2 methines and three sp^2 and one sp^3 quaternary carbons. The four double bonds accounted for four degrees of unsaturation, indicating the presence of one ring in **6**. The connectivity of proton and carbon were assigned by analysis of the ^1H–^{13}C HMQC spectrum (Supplementary Figure S38), as shown in Figure 2c. Analysis of the ^1H–^1H COSY (Supplementary Figure S39) revealed three partial structures C-1–C-2, C-5–C-9 and C-12–C-14 (Figure 2c) and the ^1H–^{13}C long-range couplings of *2J* and *3J* observed in the ^1H–^{13}C HMBC (Supplementary Figure S40) experiments gave the following information: the cross peaks from H_2-2 (δ_H 4.91, 4.89) to C-10 (δ_C 40.0), from H_2-3 (δ_H 4.81, 4.59) to C-5 (δ_C 52.8), from H_3-16 (δ_H 1.68) to C-14 (δ_C 124.2), C-15 (δ_C 131.6) and C-17 (δ_C 17.7), from H_3-17 (δ_H 1.61) to C-14, C-15 and C-16 (δ_C 25.7), from H_2-18 (δ_H 4.79, 4.72) to C-7 (δ_C 44.4), C-11 (δ_C 154.5) and C-12 (δ_C 34.9), from H_3-19 (δ_H 0.99) to C-1 (δ_C 150.1), C-5 (δ_C 52.8), C-9 (δ_C 40.1) and C-10 (δ_C 40.0), and from H_3-20 (δ_H 1.69) to C-3 (δ_C 112.0), C-4 (δ_C 147.8) and C-5 (δ_C 52.8) supported the structure as shown in Figure 2c. The stereochemical assignments of two methyl protons (H_3-19 and H_3-20) were deduced from the ^1H–^1H NOESY data (Supplementary Figure S41), whereas the C-7 prenyl residue was assigned an equatorial configuration, based on the observed correlation of H-7 with H-5 and H-9.

The HR-MS (EI) of **7** showed a molecular ion peak at *m/z* 272.2506 $[M]^+$, consistent with a molecular formula of $C_{20}H_{32}$ (calculated. 272.2504) for a diterpene hydrocarbon having five degrees of unsaturation. The ^1H NMR spectrum (Supplementary Figure S43) showed three olefinic protons (δ_H 4.51, m, 1H, H-5; 4.71, m, 1H, H-1; 5.04, m, 1H, H-14; 4.81, brs, 1H, H-3; 4.89, s, 1H, H-2; 4.91, d, $J = 17.5$ Hz, 1H, H-2; Figure 2d). The ^{13}C NMR (Supplementary Figure S43) and DEPT spectra of **7** showed 20 resolved signals and confirmed the presence of five methyls, seven sp^3 methylenes, three sp^2 methines, and five sp^2 quaternary carbons. With four double bonds the remaining degree of unsaturation was attributed to a ring in **7**. The connectivity of proton and carbon were assigned by analysis of the ^1H–^{13}C HMQC spectrum (Supplementary Figure S44), as shown in Figure 2d. Analysis of the ^1H–^1H COSY (Supplementary Figure S45) revealed four partial structures C-1–C-3, C-5–C-6, C-8–C-9, and C-12–C-14 (Figure 2d) and the ^1H–^{13}C long-range couplings of *2J* and *3J* observed in the ^1H–^{13}C HMBC (Supplementary Figure S46) experiments gave the following information: The cross peaks from H_3-16 (δ_H 1.69) to C-14 (δ_C 124.6), C-15 (δ_C 131.6) and C-17 (δ_C 17.7), from H_3-17 (δ_H 1.62) to C-14, C-15 and C-16 (δ_C 25.8), from H_3-18 (δ_H 1.68) to C-7 (δ_C 124.7), C-11 (δ_C 131.7) and C-12 (δ_C 34.7), from H_3-19 (δ_H 1.58) to C-1 (δ_C 128.4), C-9 (δ_C 40.0) and C-10 (δ_C 131.4), and from H_3-20 (δ_H 1.53) to C-3 (δ_C 39.0), C-4 (δ_C 135.8) and C-5 (δ_C 124.6) supported the structure as shown in Figure 2d.

Sesquiterpenes from transformants carrying *sclav_p1407*

Although *S. avermitilis* SUKA22 transformants carrying *sclav_p1407* produced eight sesquiterpenes, one component was produced as a major product (Supplementary Figure S3; peak 1). From the mycelial extracts obtained from 4-liters culture, the compound corresponding to peak 1 (isohirsut-1-ene; **8**, 1.7 mg, $[\alpha]_D^{24}$ -162.0 (*c* 0.05, $CHCl_3$), IR ν_{max} (attenuated total reflection) 2921, 1440, 1382 and 890 cm^{-1}) was purified by silica gel column chromatography. The HR-MS (EI) of **8** showed a molecular ion peak at *m/z* 204.1879 $[M]^+$ consistent with a sesquiterpene hydrocarbon of the molecular formula $C_{15}H_{24}$ (calculated 204.1878) with four degrees of unsaturation. The ^1H NMR spectrum (Supplementary Figure S49) showed a single olefinic proton (δ_H 5.01, s, 1H, H-1; Figure 3a). ^{13}C NMR (Supplementary Figure S49) and DEPT spectra of **8** showed 15 resolved signals and confirmed the presence of four methyls, four sp^3 methylenes, three sp^3 methines, one sp^2 methine and one sp^2 and two sp^3 quaternary carbons. In addition to the one double bond, the remaining degrees of unsaturation in **8** were attributed to three rings. The connectivity of proton and carbon were assigned by analysis of the ^1H–^{13}C HMQC spectrum (Supplementary Figure S50), as shown in Figure 3a. Analysis of the ^1H–^1H COSY (Supplementary Figure S51) revealed one partial structure C-8–C-10 (Figure 3a) and the ^1H–^{13}C long-range couplings of *2J* and *3J* observed in the ^1H–^{13}C HMBC (Supplementary Figure S52) experiments gave the following information: The cross peaks from H_1-1 (δ_H 5.01) to C-2 (δ_C 154.1) and C-9 (δ_C 47.9), from H_1-3 (δ_H 1.69) to C-1 (δ_C 127.4) and C-2, from H_2-6 (δ_H 1.50) to C-4 (δ_C 43.7), from H_2-8 (δ_H 1.82) to C-2 and C-3 (δ_C 57.7), from H_1-9 (δ_H 3.16) to C-1 and C-2, from H_2-10 (δ_H 1.76) to C-1 and C-2, from H_3-12 (δ_H 1.08) to C-1, C-10 (δ_C 47.7), C-11 (δ_C 50.8) and C-13 (δ_C 28.1), from H_3-13 (δ_H 1.02) to C-1, C-10, C-11 and C-12 (δ_C 30.1), from H_3-14 (δ_H 1.06) to C-3 (δ_C 57.7), C-5 (δ_C 35.1), C-7 (δ_C 55.3)



NMR chemical shifts in $CDCl_3$

	8		9	
No.	δ_C	δ_H (multi, *J* in Hz)	δ_C	δ_H (multi, *J* in Hz)
1	127.4	5.01 (1H, s)	48.8	1.58 (1H, dd, *J* = 13.0, 8.6), 1.30 (1H, m)
2	154.1	-	50.1	2.34 (1H, q, *J* = 10.0)
3	57.7	1.69 (1H, m)	66.7	2.11 (1H, brs)
4	43.7	*not assigned*	142.8	-
5	35.1	1.67 (1H, m), 1.20 (1H, m)	121.8	5.06 (1H, brs)
6	41.2	1.50 (2H, m)	47.5	2.19 (1H, dq, *J* = 15.7, 2.2), 2.01 (1H, dt, J = 15.7, 2.1)
7	55.3	-	52.9	-
8	47.8	1.22 (1H, m), 1.82 (1H, dd, *J* = 7.0, 11.5)	48.9	1.81(1H, dd, *J* = 13.0, 7.8), 1.32 (1H, dd, *J* = 13.0, 7.7)
9	47.9	3.16 (1H, quin, *J* = 8.3)	44.3	2.52 (1H, m)
10	47.7	0.84 (1H, m), 1.76 (1H, dd, *J* = 7.1, 11.5)	47.4	1.64 (1H, ddd, *J* = 11.6, 8.7, 1.1), 1.27 (1H, m)
11	50.8	-	41.6	-
12	30.1	1.08 (3H, s)	29.2	0.92 (3H, s)
13	28.1	1.02 (3H, s)	30.5	1.06 (3H, s)
14	30.1	1.06 (3H, s)	29.4	1.17 (3H, s)
15	20.1	1.03 (3H, d, *J* = 6.8)	15.9	1.68 (3H, s)

Figure 3 Structure of isohirsut-1-ene ((**a**); **8**) produced by *S. avermitilis* SUKA22 carrying *sclav_p1407* and isohirsut-4-ene ((**b**); **9**) produced by *S. avermitilis* SUKA22 carrying *slt18_1880*. The upper panel shows ^1H–^1H COSY (bold line), ^1H–^{13}C HMBC (arrow) and ^1H–^1H ROESY (dashed arrow) data for each compound. Double arrows indicate a cross peak in ^1H–^{13}C HMBC. A full color version of this figure is available at *The Journal of Antibiotics* journal online.

and C-8 (δ_C 47.8) and from H_3-15 (δ_H 1.03) to C-3, C-4 and C-5 supported the structure with a novel linear triquinane skeleton as shown in Figure 3a.

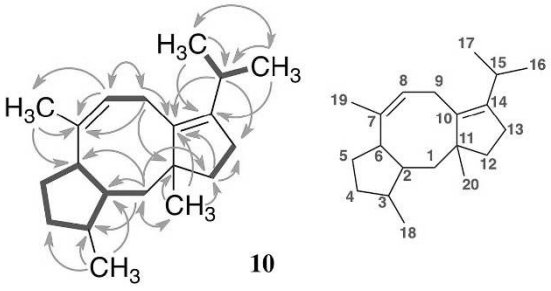
Sesquiterpenes from transformants carrying *slt18_1880*

S. avermitilis SUKA22 transformants carrying *slt18_1880* of *S. lactacystinaeus* OM-6159 produced four sesquiterpenes, with one component being the predominant metabolite (Supplementary Figure S4; peak 2). The compound corresponding to peak 2 (isohirsut-4-ene; **9**, 1.8 mg, $[\alpha]_D^{24}$ -188.0 (*c* 0.05, $CHCl_3$), IR ν_{max} (attenuated total reflection) 2921, 1440, 1382 and 890 cm^{-1}) was purified from 2 liters of culture by silica gel column chromatography. All spectroscopic data were very similar to those of **8**. The HR-MS (EI) of **9** showed a molecular ion peak at *m/z* 204.1878 $[M]^+$, consistent with a sesquiterpene hydrocarbon of molecular formula $C_{15}H_{24}$ (calculated. 204.1878) with four degrees of unsaturation. The ^1H NMR spectrum (Supplementary Figure S54) showed one olefinic proton (δ_H 5.06, brs, 1H, H-5; Figure 3b). Analysis of the ^{13}C NMR (Supplementary Figure S54) and DEPT spectra of **9** showed 15 resolved signals and confirmed the presence of four methyls, four sp^3 methylenes, three sp^3 methines, one sp^2 methine and one sp^2 and two sp^3 quaternary carbons. Besides the one double bond, the remaining three degrees of unsaturation in **9** were attributed to three rings. The connectivity of proton and carbon were assigned by analysis of the ^1H–^{13}C HMQC (Supplementary Figure S55) spectrum, as shown in Figure 3b. Analysis of the ^1H–^1H COSY (Supplementary Figure S56) revealed two partial structures C-1 to C-10 and C-5 to C-6 (Figure 3b) and the ^1H–^{13}C long-range couplings of *2J* and *3J* observed in the ^1H–^{13}C HMBC (Supplementary Figure S57) experiments gave the following information: the cross peaks from H_3-12 (δ_H 0.92) to C-1 (δ_C 48.8), C-10 (δ_C 47.4), C-11 (δ_C 41.6) and C-13 (δ_C 30.5), from H_3-13 (δ_H 1.06) to C-1, C-10, C-11 and C-12 (δ_C 29.2), from H_3-14 (δ_H 1.17) to C-6 (δ_C 47.5) C-7 (δ_C 52.9) and C-8 (δ_C 48.9) and H_3-15 (δ_H 1.68) to C3 (δ_C 66.7), C-4 (δ_C 142.8) and C-5 (δ_C 121.8) supported the structure, an isomer of **8**, as shown in Figure 3b.

Cyclooctat-7(8),10(14)-diene (10), a diterpene isolated from transformants carrying *slt18_1078*

S. avermitilis SUKA22 transformants carrying *slt18_1078* of *S. lactacystinaeus* OM-6159 produced a diterpene hydrocarbon (Supplementary Figure S5), assigned as cyclooctat-7(8),10(14)-diene, **10** (24.0 mg, $[\alpha]_D^{24}$ +70.7 (*c* 0.1, $CHCl_3$), IR ν_{max} (attenuated total reflection) 2921, 1440, 1382 and 890 cm^{-1}), which was isolated from 7 liters of culture and purified by silica gel column chromatography. The HR-MS (EI) of **10** showed a molecular ion peak at *m/z* 272.2505 $[M]^+$ consistent with a diterpene hydrocarbon of molecular formula

$C_{20}H_{32}$ (calculated. 272.2504) with five degrees of unsaturation. The ^{1}H NMR spectrum (Supplementary Figure S60) showed an olefinic proton (δ_H 5.44, dd, J = 7.7, 12.8 Hz, 1H, H-8; Figure 4). ^{13}C NMR (Supplementary Figure S60) and DEPT spectra of **10** showed 20 resolved signals and confirmed the presence of five methyls, six sp^3 methylenes, four sp^3 methines, one sp^2 methine and three sp^2 and one sp^3 quaternary carbons. Besides the two double bonds, the remaining degrees of unsaturation in **10** were attributed to three rings. The connectivity of proton and carbon were assigned by analysis of the ^{1}H–^{13}C HMQC spectrum (Supplementary Figure S61), as shown in Figure 4. Analysis of the ^{1}H–^{1}H COSY (Supplementary Figure S62) revealed four partial structures C-1–C-18, C-8–C-9, C-12–C-13 and C-15–C-17 (Figure 4) and the ^{1}H–^{13}C long-range couplings of *2J* and *3J* observed in the ^{1}H–^{13}C HMBC (Supplementary Figure S63) experiments gave the following information: The cross peaks from H_2-1 (δ_H 1.60) to C-3 (δ_C 43.0), C-6 (δ_C 49.1) and C-12 (δ_C 37.3), H_2-1 (δ_H 1.73) to C-2 (δ_C 53.5), C-6, C-12 and C-20 (δ_C 26.0), from H_1-8 (δ_H 5.44) to C-7 (δ_C 139.6), C-9 (δ_C 23.4) and C-19 (δ_C 20.6), from H_2-9 (δ_H 2.45) to C-7, C-8 (δ_C 123.8), C-10 (δ_C 141.9) and C-11 (δ_C 51.3), from H_2-12 (δ_H 1.40, 1.51) to C-10 and C-13 (δ_C 26.3), from H_2-13 (δ_H 1.99, 2.11) to C-10, from H_1-15 (δ_H 2.63) to C-10 and C-13, from H_3-16 (δ_H 0.90) to C-14 (δ_C 140.0), C-15 (δ_C 27.8) and C-17 (δ_C 21.5), from H_3-17 (δ_H 0.97) to C-14, C-15 and C-16 (δ_C 20.9), from H_3-18 (δ_H 0.93) to C-2, C-3 and C-4 (δ_C 35.9), from H_3-19 (δ_H 1.65) to C-6, C-7 and C-8 and from H_3-20 (δ_H 0.92) to C-1 (δ_C 48.2), C-10, C-11 and C-12 supported the structure as shown in Figure 4.



NMR chemical shifts in $CDCl_3$

No.	δ_C	δ_H (multi, J in Hz)
1	48.2	1.60 (1H, m), 1.73 (1H, m)
2	53.5	0.91 (1H, m)
3	43.0	1.36 (1H, m)
4	35.9	1.05 (1H, m), 1.69 (1H, m)
5	29.7	1.24 (1H, m), 1.62 (1H, m)
6	49.1	2.73 (1H, dd, J = 5.5, 8.3)
7	139.6	-
8	123.8	5.44 (1H, dd, J = 12.8, 7.7)
9	23.4	2.45 (2H, m)
10	141.9	-
11	51.3	-
12	37.3	1.40 (1H, m), 1.51 (1H, m)
13	26.3	1.99 (1H, m), 2.11 (1H, m)
14	140.0	-
15	27.8	2.63 (1H, m)
16	20.9	0.90 (3H, d, J = 7.6)
17	21.5	0.97 (3H, d, J = 7.6)
18	18.0	0.93 (3H, d, J = 7.6)
19	20.6	1.65 (3H, s)
20	26.0	0.92 (3H, s)

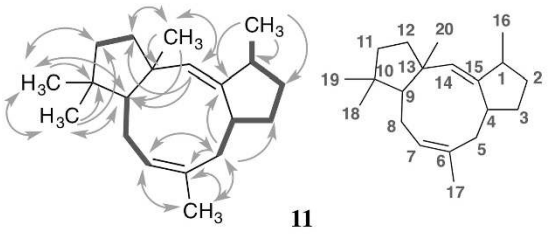
Figure 4 Structure of cyclooctat-7(8),10(14)-diene (**10**) produced by *S. avermitilis* SUKA22 carrying *slt18_1078*. The upper panel shows ^{1}H–^{1}H COSY (bold line) and ^{1}H–^{13}C HMBC (arrow) data for **10**. Double arrows indicate a cross peak in ^{1}H–^{13}C HMBC. A full color version of this figure is available at *The Journal of Antibiotics* journal online.

Diterpenes from transformants carrying *stsu_20912*

S. avermitilis SUKA22 transformants carrying *stsu_20912* of *S. tsukubaensis* NRRL 18488, the producer of the immunosuppressant polyketide tacrolimus, generated one major diterpene product along with another seven would be minor diterpene components (Supplementary Figure S6). The major component, tsukubadiene (**11**, 9.4 mg, $[\alpha]_D^{24}$ -120.3 (c 0.1, $CHCl_3$), IR ν_{max} (attenuated total reflection) 2921, 1440, 1382 and 890 cm^{-1}) was isolated from 2 liters of culture by silica gel column chromatography. The HR-MS (EI) of **11** showed a molecular ion peak at m/z 272.2506 $[M]^+$, consistent with a diterpene hydrocarbon of molecular formula $C_{20}H_{32}$ (calculated 272.2504) with five degrees of unsaturation. The ^{1}H NMR spectrum (Supplementary Figure S65) showed two olefinic protons (δ_H 5.11, t, J = 2.0, 2.0 Hz, 1H, H-14; 5.49, t, J = 7.3, 1H, H-7; Figure 5). ^{13}C NMR (Supplementary Figure S65) and DEPT spectra of **11** showed 20 resolved signals and confirmed the presence of five methyls, six sp^3 methylenes, three sp^3 methines, two sp^2 methines and two sp^2 and two sp^3 quaternary carbons. Besides the two double bonds, the remaining three degrees of unsaturation were attributed to three rings in **11**. The connectivity of proton and carbon were assigned by analysis of the ^{1}H–^{13}C HMQC spectrum (Supplementary Figure S66), as shown in Figure 5. Analysis of the ^{1}H–^{1}H COSY (Supplementary Figure S67) revealed three partial structures C-1–C-16, C-7–C-9, C-12 and C-11–C-12 (Figure 5) and the ^{1}H–^{13}C long-range couplings of *2J* and *3J* observed in the ^{1}H–^{13}C HMBC (Supplementary Figure S68) experiments gave the following information: The cross peaks from H_2-5 (δ_H 2.25, 2.01) to C-3 (δ_C 37.0), C-6 (δ_C 137.5), C-7 (δ_C 124.7), C-15 (δ_C 143.7) and C-17 (δ_C 25.3), from H_1-7 (δ_H 5.49) to C-5 (δ_C 41.3) and C-17, from H_1-9 (δ_H 2.34) to C-18 (δ_C 32.1) and C-19 (δ_C 25.7), from H_2-11 (δ_H 1.48, 1.37) to C-19, from H_2-12 (δ_H 1.47) to C-20 (δ_C 25.4), from H_1-14 (δ_H 5.11) to C-9 (δ_C 53.3), C-12 (δ_C 40.7) and C-20, from H_3-16 (δ_H 1.05) to C-1 (δ_C 41.7), C-2 (δ_C 31.9) and C-15 (δ_C 143.7), from H_3-18 (δ_H 1.00) to C-9, C-10 (δ_C 40.8), C-11 (δ_C 39.9) and C-19, from H_3-19 (δ_H 0.93) to C-9, C-10 and C-18, and from H_3-20 (δ_H 0.96) to C-9 (δ_C 53.3), C-12, C-13 (δ_C 48.1) and C-14 (δ_C 132.4) supported the structure with a novel 5-9-5 ring skeleton as shown in Figure 5.

Diterpenes from transformants carrying *nd90_0354*

S. avermitilis SUKA22 transformants carrying *nd90_0354* of *Streptomyces* sp. ND90 produced two major and one minor diterpene components (Supplementary Figure S7). The two major components were isolated from 2 liters culture, with odyverdienes A (**12**, 1.2 mg, $[\alpha]_D^{24}$ -18.7 (c 0.1, $CHCl_3$), IR ν_{max} (attenuated total reflection) 2921, 1440, 1382, 890 cm^{-1}) and B (**13**, 1.2 mg, $[\alpha]_D^{24}$ -44.2 (c 0.1, $CHCl_3$), IR ν_{max} (attenuated total reflection) 2921, 1440, 1382, 890 cm^{-1}) each being obtained as a colorless oil. The HR-MS (EI) of **12** showed a molecular ion peak at m/z 272.2506 $[M]^+$, corresponding to a diterpene hydrocarbon of molecular formula $C_{20}H_{32}$ (calculated 272.2504) with five degrees of unsaturation. The ^{1}H NMR spectrum (Supplementary Figure S71) showed four olefinic protons (δ_H 4.63, s, 1H, H-17; 4.66, s, 1H, H-17; 4.85, s, 1H, H-18; 4.96, s, 1H, H-18; Figure 6a) assignable to two exomethylenes. ^{13}C NMR (Supplementary Figure S71) and DEPT spectra of **12** showed 20 resolved signals and confirmed the presence of three methyls, seven sp^3



NMR chemical shifts in $CDCl_3$

No.	δ_C	δ_H (multi, *J* in Hz)
1	41.7	2.32 (1H, m)
2	31.9	1.70 (1H, m), 1.13 (1H, m)
3	37.0	1.78 (1H, m), 1.58 (1H, m)
4	40.9	2.96 (1H, m)
5	41.3	2.25 (1H, dd, *J* = 13.5, 7.7), 2.01 (1H, dd, *J* = 13.5. 2.8)
6	137.5	-
7	124.7	5.49 (1H, t, *J* = 7.3)
8	24.0	1.90 (2H, m)
9	53.3	2.34 (1H, t, *J* = 7.0)
10	40.8	-
11	39.9	1.48 (1H, m), 1.37 (1H, m)
12	40.7	1.47 (2H, m)
13	48.1	-
14	132.4	5.11 (1H, t, *J* = 2.0)
15	143.7	-
16	21.7	1.05 (3H, d, *J* = 7.0)
17	25.3	1.64 (3H, s)
18	32.1	1.00 (3H, s)
19	25.7	0.93 (3H, s)
20	25.4	0.96 (3H, s)

Figure 5 Structure of tsukubadiene (**11**) produced by *S. avermitilis* SUKA22 carrying *stsu_20912*. The upper panel shows ^1H–^1H COSY (bold line) and ^1H–^{13}C HMBC (arrow) data for **11**. Double arrows indicate a cross peak in ^1H–^{13}C HMBC. A full color version of this figure is available at *The Journal of Antibiotics* journal online.

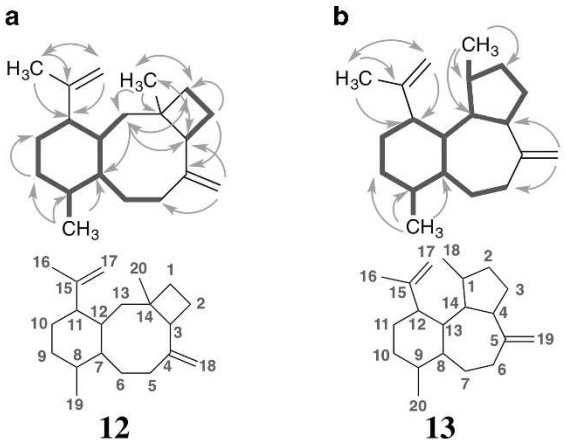
methylenes, two sp^2 methylenes, five sp^3 methines, and two sp^2 and one sp^3 quaternary carbons. Besides the two double bonds the remaining degrees of unsaturation in **12** were attributed to three rings. The connectivity of proton and carbon were assigned by analysis of the ^1H-^{13}C HMQC spectrum (Supplementary Figure S72), as shown in Figure 6a. Analysis of the ^1H–^1H COSY (Supplementary Fig. S73) revealed two partial structures C-1–C-3 and C-5–C-19 (Figure 6a) and the ^1H–^{13}C long-range couplings of *2J* and *3J* observed in the ^1H–^{13}C HMBC (Supplementary Figure S74) experiments gave the following information: the cross peaks from H_2-1 (δ_H 1.57) to C-13 (δ_C 44.9) and C-20 (δ_C 26.8), from H_2-2 (δ_H 2.07) to C-1 (δ_C 33.8) and C-4 (δ_C 150.5), from H_1-3 (δ_H 2.36) to C-20, from H_2-9 (δ_H 1.62) to C-10 (δ_C 32.4), from H_2-13 (δ_H 1.36) to C-1, C-3 (δ_C 47.8) and C-7 (δ_C 49.9), from H_3-16 (δ_H 1.58) to C-11 (δ_C 52.9), C-15 (δ_C 150.3) and C-17 (δ_C 110.8), from H_2-17 (δ_H 4.66, 4.63) to C-11 and C-16 (δ_C 19.1), from H_2-18 (δ_H 4.96, 4.85) to C-3 and C-5 (δ_C 40.6), from H_3-19 (δ_H 0.90) to C-7, C-8 (δ_C 39.2) and C-9 (δ_C 35.1) and from H_3-20 (δ_H 1.29) to C-1, C-3, C-13 (δ_C 44.9) and C-14 (δ_C 41.2) supported the structure with a novel 6-8-4 ring skeleton as shown in Figure 6a.

The HR-MS (EI) of **13** also showed a molecular ion peak at *m/z* 272.2506 $[M]^+$, corresponding to a diterpene hydrocarbon of molecular formula $C_{20}H_{32}$ (calculated 272.2504) with five degrees of unsaturation. The ^1H NMR spectrum (Supplementary Figure S76) showed four olefinic protons (δ_H 4.58, s, 1H, H-17; 4.70, s, 1H, H-17; 4.75, s, 1H, H-19; 4.89, s, 1H, H-19; Figure 6b) assignable to two exomethylenes. Analysis of ^{13}C NMR (Supplementary Fig. S76) and DEPT spectra of **13** showed 20 resolved signals and confirmed the presence of three methyls, six sp^3 methylenes, two sp^2 methylenes, seven sp^3 methines and two sp^2 quaternary carbons. Besides the two double bonds the remaining degrees of unsaturation in **12** were attributed to three rings in **13**. The connectivity of proton and carbon were assigned by analysis of the ^1H–^{13}C HMQC (Supplementary Figure S77) spectrum, as shown in Figure 6b. Analysis of the ^1H–^1H COSY (Supplementary Figure S78) revealed one partial structure C-1 to C-20 (Figure 6b) and the ^1H–^{13}C long-range couplings of *2J* and *3J* observed in the ^1H–^{13}C HMBC (Supplementary Figure S79) experiments gave the following information: The cross peaks from H_3-16 (δ_H 1.64) to C-12 (δ_C 56.0), C-15 (δ_C 150.7) and C-17 (δ_C 110.8), from H_2-17 (δ_H 4.70, 4.58) to C-12 and C-16 (δ_C 20.1), from H_3-18 (δ_H 0.83) to C-1 (δ_C 35.1), C-2 (δ_C 29.5) and C-14 (δ_C 59.6), H_2-19 (δ_H 4.89, 4.75) to C-4 (δ_C 42.4) and C-6 (δ_C 38.1) and from H_3-20 (δ_H 0.90) to C-8 (δ_C 46.0), C-9 (δ_C 37.5) and C-10 (δ_C 35.4) supported the structure with a novel 6-7-5 ring skeleton as shown in Figure 6b.

DISCUSSION

Actinomycetales microorganisms are prolific producers of an enormous variety of natural products, including aminoglycosides, polyketides, peptides, shikimate-derived aromatic compounds, among many others. By contrast, reports describing the production of terpenoid metabolites by these microorganisms have been very limited to date. Genome sequencing has established that *Actinomycetales* microorganisms harbor many genes, many within extensive gene clusters, dedicated to secondary metabolite biosynthesis. Our recent studies have revealed that genes encoding presumptive terpene synthase are in fact widely distributed in bacteria.[9] *Actinomycetales* microorganisms are especially rich in genes encoding terpene synthases, although the majority of such genes appear to be cryptic in the parent microorganisms when studied under standard laboratory culture conditions. To assign the biochemical function of these apparently silent terpene synthases, we have therefore used heterologous expression to evaluate the intrinsic ability of these bacterial genes to control the biosynthesis of terpenes.[9]

The closely related diterpene hydrocarbon and alcohol, hydropyrene (**1**) and hydropyrenol (**2**), were efficiently produced in a heterologous *Streptomyces* host carrying *sclav_p0756*, although they were not produced by the parent microorganism *S. clavuligerus* ATCC 26074 under any culture conditions. The unique structures of **1** and **2** have not previously been reported, with characterized by a skeleton consisting of four fused cyclohexane rings. The same heterologous transformants also produced the known diterpene hydrocarbon, isoelisabethatriene (which has previously been isolated from sea plumes), and its isomer, the novel diterpene hydrocarbon, isoelisabethatriene B (**3**) as minor components. Each of these minor components may be generated by alternative cyclization of the allylic cation formed upon initial ionization of geranylgeranyl diphosphate. The actual cyclization mechanism leading to these unusual diterpenes should be amenable to evaluation by NMR analysis of the corresponding ^{13}C-labeled compounds obtained by feeding ^{13}C-labeled precursors to the *S. avermitilis* host carrying *sclav_p0756*. It is predicted that the six diterpenes, **4**, **5**, **6** and **7** as well as prenylgermacrene and lobophytumin C, which were produced by transformants carrying



NMR chemical shifts in $CDCl_3$

	12		13	
No.	δ_C	δ_H (multi, *J* in Hz)	δ_C	δ_H (multi, *J* in Hz)
1	33.8	1.57 (2H, m)	35.1	2.29 (1H, m)
2	16.9	2.07 (1H, m), 1.79 (1H, m)	29.5	1.80(1H, m), 1.24 (1H, m)
3	47.8	2.36 (1H, dd, *J* = 9.8, 3.5)	27.0	1.58 (2H, m)
4	150.5	-	42.4	2.52 (1H, m)
5	40.6	2.17 (1H, m), 1.96 (1H, ddd, *J* = 14.0, 14.0, 3.0) 1.53 (1H, m),	151.6	- 2.36 (1H, m),
6	31.0	1.22 (1H, m)	38.1	2.11 (1H, ddd, *J* = 13.4, 13.4, 5.8)
7	49.9	0.42 (1H, m)	28.3	1.70 (1H, m), 1.44 (1H, m)
8	39.2	1.06 (1H, m)	46.0	0.75 (1H, m)
9	35.1	1.62 (1H, m), 0.96 (1H, m)	37.5	1.14 (1H, m)
10	32.4	1.49 (1H, m), 1.29 (1H, m)	35.4	1.67 (1H, m), 0.98 (1H, m)
11	52.9	1.55 (1H, m)	32.9	1.43 (1H, m), 1.35 (1H, m)
12	41.1	1.10 (1H, m)	56.0	1.64 (1H, m)
13	44.9	1.36 (1H, m), 1.14 (1H, m)	45.7	1.28 (1H, m)
14	41.2	-	59.6	1.11 (1H, m)
15	150.3	-	150.7	-
16	19.1	1.58 (3H, s)	20.1	1.64 (3H, s)
17	110.8	4.66 (1H, s), 4.63 (1H, s)	107.0	4.70 (1H, s), 4.58 (1H, s)
18	108.1	4.96 (1H, s), 4.85 (1H, s)	19.8	0.83 (3H, d, *J* = 12.0)
19	20.4	0.90 (3H, d, *J* = 12.0)	110.0	4.89 (1H, s), 4.75 (1H, s)
20	26.8	1.29 (3H, s)	20.6	0.90 (3H, d, *J* = 12.0)

Figure 6 Structures of odyverdienes (**a**) (A; **12**) and (**b**) (B; **13**) produced by *S. avermitilis* SUKA22 carrying *nd90_0354*. The upper panel shows ^1H–^1H COSY (bold line) and ^1H-^{13}C HMBC (arrow) data for each compound. Double arrows indicate a cross peak in ^1H-^{13}C HMBC. A full color version of this figure is available at *The Journal of Antibiotics* journal online.

sclav_p1169, are likely generated by a common cyclization mechanism involving the prenylgermacrene cation.

The production of sesquiterpenes possessing a triquinane skeleton by bacteria has not previously been reported. Interestingly, two triquinane-type sesquiterpene hydrocarbons, isohirsut-1-ene (**8**) and isohirsut-4-ene (**9**), were generated by different terpene synthases, SCLAV_p1407 and SLT18_1880, respectively which exhibited limited amino acid sequence similarity, with SLT18_1880 (325 aa) showing 32% identity and 47% similarity to SCLAV_p1407 (367 aa). SLT18_1880 was also similar to a monoterpene synthase SCLAV_p0982 (31% identity and 48% similarity), which catalyzes the generation of 1,8-cineole, camphene and β-pinene.[9]

Metabolites with a fused three ring 5-9-5, 6-8-4 and 6-7-5 skeletons have not previously been reported. In these studies, we have described the production and structure elucidation of three such unusual diterpene hydrocarbons, tsukubadiene (**11**) and odyverdines A (**12**) and B (**13**), The 6-8-4 ring system of odyverdiene A (**12**) differs from the 6-7-5 fused rings of odyverdiene B (**13**), representing alternative modes of cyclization of their common geranylgeranyl diphosphate precursor.

Given the wealth of new terpenes uncovered in this initial survey, it is evident that bacterial terpene synthases have the potential to generate unique structures and that the genes encoding these terpene synthase are a very-attractive source for the discovery of new natural products.

MATERIALS AND METHODS

Bacterial strains and growth conditions

The previously described large-deletion derivative of *S. avermitilis*, SUKA22,[10] was used as the heterologous host for the expression of genes encoding presumptive terpene synthases derived from *Streptomyces* microorganisms. Cloning of genes encoding terpene synthase and the cultivation conditions for the heterologous expression of terpene synthase genes were described previously.[10,11] Two integrating vectors, pKU1021*fps* (AB982125), also harboring a gene for farnesyl diphosphate synthase, and pKU1021*ggs* (AB982126), harboring the gene for geranylgeranyl diphosphate synthase, were used for expression in *S. avermitilis* SUKA22 of genes encoding sesquiterpene synthases

and diterpene synthases, respectively, *S. avermitilis* SUKA22 carrying pKU460[11] served as the negative control.

Isolation of terpenoid metabolites

Terpene hydrocarbons and alcohols were extracted with methanol from the mycelium of *S. avermitilis* SUKA22 transformants carrying individual genes encoding heterologous terpene synthases. The methanol extract was re-extracted with hexane and a portion of the hexane extract was directly analyzed by GC-MS using previously described analytical conditions.[12] Terpenes for which there was no match in the current GC-MS databases, NIST/EPA/NIH MS Library (2014 version), were purified by silica gel chromatography, as described previously.[9]

Physico-chemical analysis

Nuclear magnetic resonance (NMR) spectra were obtained on a JEOL JNM-ECP 500 FT NMR System (^{1}H: 500 MHz, ^{13}C: 125 MHz) and Agilent Technologies UNITY-400 (^{1}H: 400 MHz). Chemical shifts are referenced to $CDCl_3$ at room temperature. HR-MS spectra by electron ionization (EI) mode were obtained on a JEOL JMS-700 Mstation. Infrared (IR) spectra were recorded on a Horiba FT-720 infrared spectrometer and optical rotations were recorded on a Horiba SEPA-300 polarimeter.

Preparation of hydropyrene 4,12-epoxide

Purified hydropyrene (5.2 mg, 0.019 mmol) was dissolved in 1 ml of CH_2Cl_2 under N_2 atmosphere at 0 °C. *m*-Chloroperbenzoic acid (4.9 mg, 0.023 mmol) was added and the mixture was stirred for 1 h at 0 °C. After the reaction was stopped by addition of 1 ml $Na_2S_2O_3$-saturated water, the product was extracted twice with 5 ml of CH_2Cl_2. The organic layers were combined, dried over anhydrous Na_2SO_4, and evaporated to dryness. The residue was dissolved in small volume of hexane and the oxidized product was separated by silica gel column chromatography. The purified product eluted with hexane/ethyl acetate (50:1) to give 4.1 mg of hydropyrene-4,12-epoxide, which was obtained as white powder (yield 79%; HR-MS (EI) $[M]^+$ calculated for $C_{20}H_{32}O$ 288.2453, found 288.2455).

X-ray crystallographic analysis

Hydropyrene-4,12-epoxide (3.5 mg) was crystallized from chloroform. The colorless crystal having approximate dimension of $0.18 \times 0.14 \times 0.10$ mm was mounted. The single crystal X-ray diffraction data were collected on a Rigaku R-AXIS Varimax RAPID-II with IP area detectors system using Cu-Kα radiation ($\lambda = 1.54178$ Å). The crystal structure was solved by the direct method by *SHELXS*.[13] Refinement was performed by full-matrix least-squares using *SHELXL-2014*. Crystallographic data for the structure of hydropyrene-4,12-epoxide have been deposited with the Cambridge Crystallographic Data Center, deposition No. CCDC-1022903.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work supported by a research Grant-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to HI), a research Grant-in-Aid for 'Project focused on developing key technology of discovering and manufacturing drug for next-generation treatment and diagnosis' from the Ministry of Economy, Trade and Industry of Japan (to KS and HI), and a grant from the US National Institutes of General Medical Sciences, GM30301 (to DEC).

1 Berthelot, M. & André, G. Sur l'odeur proper de la terre. *Compt. Rend. Acad. Sci.* **112**, 598–599 (1891).
2 Gerber, N. N Geosmin, from microorganisms, is trans-1,10-dimethyl-trans-9-decalol. *Tetrahedron Lett.* **9**, 2971–2974 (1968).
3 Gerber, N. N. A volatile metabolite of actinomycetes, 2-methylisoborneol. *J. Antibiot.* **22**, 508–509 (1969).
4 Izaguirre, G., Hwang, C. J., Krasner, S. W. & McGuire, M. J. Geosmin and 2-methylisoborneol from cyanobacteria in three water-supply systems. *Appl. Environ. Microbiol.* **43**, 708–714 (1982).
5 Jüttner, F. Volatile organic substances. *The Cyanobacteria* (eds Fay B., Baalen C. V.) 453–469 (Elsevier, Amsterdam, 1987).
6 Wu, J-T. & Jüttner, F. Differential partitioning of geosmin and 2-methylisoborneol between cellular constituents in *Oscillatoria tenuis*. *Arch. Microbiol.* **150**, 580–583 (1988).
7 Komatsu, M., Tsuda, M., Omura, S., Oikawa, H. & Ikeda, H. Identification and functional analysis of genes controlling biosynthesis of 2-methylisoborneol. *Proc. Natl. Acad. Sci. USA* **105**, 7422–7427 (2008).
8 Yamada, Y., Cane, D. E. & Ikeda, H. Diversity and Analysis of Bacterial Terpene Synthases. In Chapter seven, Natural Product Biosynthesis by Microorganisms and Plants, Part A. *Methods in Enzymology* Volume 515 (ed. Hopwood D. A.) 123–166 (Elsevier Inc. Academic Press, New York, 2012).
9 Yamada, Y. *et al*. Terpene synthases are widely distributed in bacteria. *Proc. Natl. Acad. Sci. USA* (e-pub ahead of print 22 December 2014; doi:10.1073/pnas.1422108112).
10 Komatsu, M. *et al*. Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. *ACS Synth. Biol.* **2**, 384–396 (2013).
11 Komatsu, M., Uchiyama, T., Omura, S., Cane, D.E. & Ikeda, H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc. Natl. Acad. Sci. USA* **107**, 2646–2651 (2010).
12 Cane, D. E., He, X., Kobayashi, S., Omura, S. & Ikeda, H. Geosmin biosynthesis in *Streptomyces avermitilis*. Molecular cloning, expression, and mechanistic study of the germacradienol/geosmin synthase. *J. Antibiot.* **59**, 471–479 (2006).
13 Sheldrick, G. M. A short history of *SHELX*. *Acta Crystallogr.* **A64**, 112–122 (2008).
14 Kohl, A. C. & Kerr, R. G. Pseudopterosin biosynthesis: Aromatization of the diterpene cyclase product, elisabethatriene. *Mar. Drugs* **1**, 54–65 (2003).
15 Jakupovic, J. *et al*. Twenty-one acylphloroglucinol derivatives and further constituents from south african *Helichrysum* species. *Phytochemistry* **28**, 1119–1131 (1989).
16 Li, L. *et al*. Diterpenes from the Hainan soft coral *Lobophytum cristatum* Tixier-Durivault. *J. Nat. Prod.* **74**, 2089–2094 (2011).

Supplementary Information accompanies the paper on The Journal of Antibiotics website (http://www.nature.com/ja)